

RANDOM DIGIT DIALING IN SELECTING A POPULATION-BASED CONTROL GROUP

PATRICIA HARTGE,¹ LOUISE A. BRINTON,¹ JEANNE F. ROSENTHAL,² JOHN I. CAHILL,²
ROBERT N. HOOVER¹ AND JOSEPH WAKSBERG²

Hartge, P. (NCI, Bethesda, MD 20205), L. A. Brinton, J. F. Rosenthal, J. I. Cahill, R. N. Hoover and J. Waksberg. Random digit dialing in selecting a population-based control group. *Am J Epidemiol* 1984;120:825-33.

Results are described from four epidemiologic studies in the United States which used random digit dialing in over 30,000 households to identify controls from the general population for use in case-control studies. Methods and problems in telephone sampling are discussed. It is concluded that if complete population rosters are unavailable and if the population to be sampled has the high rates of telephone ownership typical of much of the United States, telephone-based sampling can yield a nearly random sample of the individuals in a population, often at much less expense than can dwelling-based sampling.

interviews; sampling studies; telephone; epidemiologic methods

Five years ago, we and our colleagues at the National Cancer Institute began using random digit dialing of telephone numbers as a method of selecting study subjects from the general population. This method had not been widely used for epidemiologic research before, but had been used in survey and market research (1). A population-based sample is particularly useful in case-control studies in which the case series represents cases in the entire population.

Random digit dialing is one method of selecting telephone numbers at random from the pool of telephone numbers that belong to households in a defined geographic area. Individuals in telephone-owning households can then be enumerated and sampled. The sample of individuals produced by this procedure is useful to the extent that it approximates a probability sample of the individuals in the population. This occurs if the vast majority of individuals live in households with telephones (coverage) and if people contacted by telephone or their household members are willing to be in the study (response). For many, but not all, subpopulations within the United States, telephone coverage is almost universal, and, in our experience, acceptably high response rates can be achieved in most subpopulations. In this paper, we describe why and how we have used this sampling method and present results of its use in four studies (2-5). We consider how random digit dialing compares with other methods of sampling from the general population.

Received for publication July 6, 1983, and in final form March 13, 1984.

¹ Environmental Epidemiology Branch, National Cancer Institute, Bethesda, MD.

² Westat, Inc., Rockville, MD.

Reprint requests to Dr. Patricia Hartge, Environmental Studies Section, Environmental Epidemiology Branch, National Cancer Institute, Landow Building, Room 3C09, Bethesda, MD 20205.

Dr. Arlene Kantor directed the T-cell lymphoma study, and Edward P. Ward supervised the telephone sampling. Dr. Elaine Ron designed the thyroid cancer study, Ruth Kleiner directed it, and Dan Ames supervised the telephone sampling. The authors thank them and their other colleagues for helpful comments on this manuscript and for permitting them to use their data.

METHODS

Telephone coverage

Before we decide whether to use random digit dialing to draw a population-based sample in a study, we evaluate the prevalence of telephone ownership in the population to be studied. As table 1 shows, telephone coverage is high in all regions of the United States and in most demographic and economic groups. Even in families with less than \$3,000 income, over 70 per cent of households have telephones. Ninety-four per cent of whites and 84 per cent of blacks live in households that own telephones.

If telephone ownership is relatively low in a study population, telephone-based sampling is inappropriate. Under such circumstances, two control groups may be used, a telephone-based sample to be compared with the cases who own telephones and a sample from households without telephones for other cases. In effect, the control group would be a sample stratified according to telephone ownership. We are not aware of studies using this more complex procedure.

Sampling telephone households

Telephone directories are very poor sampling frames because they omit unlisted numbers and households that have moved since the publication date of the directory. These two groups are a substantial portion of all households in most areas.

Occasionally, lists of all residential phone numbers (listed and unlisted) can be obtained from the telephone company, from which a random sample can be selected. When no such list exists, simple random sampling of households that have telephones (1) begins with obtaining a recent list of all telephone area codes and three-digit prefix numbers within the designated area. This list is sampled at random and four randomly chosen digits are appended to produce complete tele-

TABLE 1
*Household telephone ownership,
United States, 1981**

Subpopulation	% living in households with telephones
Race	
White	94.3
Black	84.5
Other	93.0
Sex	
Male	92.7
Female	93.7
Age (years)	
<5	87.4
5-14	92.3
15-24	90.2
25-34	92.8
35-44	95.0
45-54	96.2
55-64	96.5
65-74	96.8
≥75	96.4
Region	
Northeast	94.1
North central	94.9
South	90.1
West	94.2
Urban-rural residence	
Urban	93.7
Rural farm	95.3
Rural nonfarm	91.2
Family income (dollars)	
<3,000	71.6
3,000-4,999	80.4
5,000-6,999	84.2
7,000-9,999	85.4
10,000-14,999	92.3
15,000-24,999	96.4
25,000 +	99.2
Health status	
Excellent	94.8
Good	92.2
Fair	90.9
Poor	89.0
Total United States	93.2

* Source: National Center for Health Statistics, unpublished data.

phone numbers. The telephone numbers are called to determine which belong to residences. The procedure is straightforward and unbiased but time-consuming

because only about 20 per cent of telephone numbers so generated will belong to residences (1).

The Waksberg method

In 1978, Waksberg (6) showed that a clustered random sample of households with telephones, a method first used by Mitofsky (7), was unbiased and could usually be an effective means of decreasing the ratio of telephone numbers to residences. The Waksberg procedure defines as a primary sampling unit the 100 10-digit telephone numbers (area code + prefix + four-digit suffix) that share the same first eight digits. One such primary sampling unit is drawn at random; then one telephone number within that primary sampling unit is randomly selected and dialed. If the telephone number is a working residential number (the probability of which depends on the fraction of such telephones within the primary sampling unit), then that primary sampling unit is sampled; otherwise, it is rejected. Within each sampled primary sampling unit, a predetermined number of residential telephones, K , is sampled.

The probability of selection of a primary sampling unit is proportional to the fraction of residential telephones within that primary sampling unit, and the chance of a household being selected within a primary sampling unit varies as the inverse of this fraction. The overall effect is that all residential telephone numbers have the same probability of selection.

Several features of the Waksberg procedure are frequently misunderstood. First, even though primary sampling units with a small number of residential numbers are rejected more frequently than larger primary sampling units, this does not mean that households in such primary sampling units are underrepresented in the sample. The analysis of the mathematics underlying the Waksberg procedure demonstrates that all house-

holds have exactly the same probability of selection when all clusters have at least K households (K should be kept small enough so that this applies; keeping K under 35 or 40 will generally accomplish it). An example of how the sample operates in small populations may help clarify the situation.

Assume a population with clusters of 10, 20, 40, and 80 households. Let there be 20 clusters of each type. The 20 clusters of 10 households each have a total of 200 households; the 20 clusters of 20 have a total of 400 households, etc. Table 2 illustrates what happens at the two stages of selection, first, choosing clusters with equal probability and, second, rejecting clusters in which the sampled number is not residential. It can be seen that the final, expected sample is 2 per cent in all four groups.

Some comments on the choice of the value of K may be helpful. As long as the same number of households is chosen per primary sampling unit, all residential telephones in the survey area will have the same chance of selection, and the sample will comprise an unbiased, equal probability sample of residential telephones. In choosing the value of K for a particular study, an important consideration is that the time and cost of screening is affected by the value of K , and the greater the K , the lower the cost. On the other hand, to the extent homogeneity exists within clusters, higher values of K tend to reduce the power of statistical tests and thus require greater sample sizes to obtain the same power. In addition, the statistical tests generally used assume independence in the sample selection, and large values of K tend to increase dependence.

Clearly, some compromises are necessary in the choice of K . The policy we have usually followed is to choose a value which provides an average of only about one to two controls per cluster. The intra-class correlations within clusters then

TABLE 2

Identical number of clusters in each of four size groups with selection of one-half clusters in initial draw

Item	No. of households in cluster			
	10	20	40	80
No. of clusters in population	20	20	20	20
Total households in population	200	400	800	1,600
No. of clusters expected in first draw	10	10	10	10
No. rejected	9	8	6	2
No. remaining after rejected nonhouseholds	1	2	4	8
Total households expected in sample (using clusters of four households)	4	8	16	32
% of households in sample	4/200 = 2%	8/400 = 2%	16/800 = 2%	32/1,600 = 2%

have only a trivial effect on the power of the statistical tests, and the assumptions underlying the tests are quite closely met. For most case-control studies, the screening ratio is about 8 or 10 to 1 (or even greater), i.e., on average, 8 or 10 households have to be screened to identify a single control. Therefore, K is generally in the range of at least 8–20. This value is increased when it appears that the study will mainly analyze subgroups, for example, statistics on males and females. In such cases, the value of K can be doubled (assuming about the same number of controls for males and females) since then there will still be only one to two controls per relevant population group being analyzed.

Steps in sampling telephone households by the Waksberg method

Random digit dialing using the Waksberg method (6) requires estimating the sample size, with the following steps:

1) Set the geographic boundaries. Within these boundaries, identify all of the area codes and three-digit exchanges assigned by the telephone companies.

2) Based on the stratification criteria and the total study size, determine the number of subjects (M) needed within each sampling stratum.

3) From Census Bureau data or other

sources, estimate the ratio of households to study subjects in each stratum (L) (e.g., 15 households per white man aged 60–64 years in Iowa).

4) Within each stratum, multiply M by L to find the expected number of households (N) needed to yield the desired number of study subjects in that stratum.

5) Determine which stratum will require the largest number of households. This is the stratum which “drives” the sample size: a sample large enough for the driving stratum will fill the other strata.

6) Depending on the study hypotheses, determine what degree of clustering of households within a primary sampling unit will be tolerable.

7) Based on the tolerable level of intra-primary sampling unit clustering (as reflected in the choice of K) and the number of households needed in the “driving” stratum (N), determine the expected number of residential primary sampling units needed (N/K).

The procedures for calling the selected telephone numbers are then as follows:

8) Randomly select an area code, an exchange in the geographic area, and four random digits. Call the number and determine whether the telephone belongs to a residence.

9) If the number is residential, the

household members are eligible to be sampled, and the number defines a residential primary sampling unit, i.e., the 100 telephone numbers that begin with the first five digits of the telephone number called.

10) Call $K - 1$ other randomly chosen numbers within the residential primary sampling unit. If one of them is not residential, select another until the preselected number of residences (K) has been reached.

11) Follow steps 8–10 until the needed number of households of individuals has been obtained (N/K times, on the average).

*Stratified sampling of individuals
within households*

Epidemiologic studies often require a sample of subjects stratified on geography, age, race, sex, or other factors. Telephone sampling can accommodate such stratified sampling in different ways. For instance, the telephone interviewer can obtain a listing of all the household members, along with the needed data on age, sex, race, etc., from each of the K households within the primary sampling unit, and the stratified sample can be drawn later and the sampled subjects contacted by telephone, letter, or in person.

Alternatively, the telephone interviewer can apply the stratum-specific sampling probability to each household member and determine which household members fall into the sample. The interviewer can then attempt to arrange interviews with the sampled members (8). (This one-call procedure for enumerating and sampling is only feasible if the stratified sampling scheme is not complicated.)

In either stratified or unstratified sampling, an individual's probability of selection should not depend on how frequently members of his or her household are at home. Such a bias may be inadvertently

introduced by certain calling procedures. For example, if three residential households are being selected within each primary sampling unit ($K = 3$), then interviewers can legitimately try to reach as many as three telephone numbers in the primary sampling unit at one time. If so, the same stratum-specific sampling fractions must be applied to all individuals eventually identified at those three telephone numbers. A rule such as "take the first two women aged 21–24 years and stop looking thereafter" would be biased, since women aged 21–24 years at the hardest-to-reach telephone numbers would have a lowered probability of selection simply because their households did not often have someone at home. An unbiased procedure would be to complete all three telephone numbers, applying the same sampling fractions, and then to draw two women aged 21–24 years at random from those identified.

An issue of intraclass correlation also affects the sampling procedures for selecting persons within households. We have generally felt it undesirable to select more than one control per household, mostly because of our concern about the effect of intraclass correlation within households, but also to limit respondent burden. To choose one person at random within each selected household would introduce variation in chance of selection, i.e., the smaller the household, the greater would be the probability of selection. Our usual practice has been to designate households on a random basis as corresponding to either male or female controls. Only males in a male-designated household are then eligible to be controls, and similarly for females. The system is unbiased and produces an equal-probability selection of controls. This does not completely eliminate the possibility that a household will have two controls selected, but this occurrence is rare. It should be noted that the designation of households as male or female does not

have to be done on a 50-50 basis. The proportion of households assigned to each sex can be allocated on the basis of the number of controls required for each sex.

Various problems can arise in stratified sampling. Sometimes, the person answering the telephone is willing to give the information needed to determine whether any household members should be sampled, but not willing to participate fully, e.g., by providing an address in those studies where a letter is sent to all subjects selected. We consider these contacts as partially complete. If a household member is drawn into the sample, we call the household again and attempt to obtain the address. If the person answering refuses, the selected subject is considered a nonrespondent.

Maintaining the response rate

The estimated overall response rate depends both on the response rate at the stage of identifying households and enumerating their members as well as on the response rate when the individuals selected as study subjects are approached for interview. If the probabilities of non-response at the two stages were independent, the overall response rate would be the product of the two component rates.

Because nonresponse can occur at two stages (sampling households and interviewing individuals), it is as vital to achieve a high response rate during household sampling as during interviewing. Achieving a high response rate depends critically on the telephone interviewing personnel. Personnel suited to telephone interviewing share characteristics with good personal interviewers, but voice is more important and physical appearance or handicap is unimportant. In our experience, response rates vary from 60–80 per cent during the training period. Between 1979 and 1982, we found that acceptable interviewers could maintain a response rate of 85 per cent by the end of a two-week training period. Our very best interviewers could maintain a

consistent rate in the range 88–92 per cent. (During the last year, these target rates have fallen slightly, mainly in urban areas.) We often have the best interviewers assist other interviewers by calling those households where hesitation short of an outright refusal has been encountered in the first telephone call. In our experience, female interviewers have achieved a higher response rate by telephone than male interviewers.

Problems and issues

A variety of practical considerations arise when implementing the basic procedure outlined above. For example, some telephones will not be answered despite repeated calls. Interviewers can vary the times of day and days of the week to increase the chance of having the telephone answered. After seven to nine attempts have been made, with calls made on weekends and weekdays, mornings, afternoons, and evenings, we generally stop trying to reach a number. We then ask the telephone company whether the number is a working number and whether it is a residence. If the telephone is residential, we continue calling over a period of two to three weeks. Sometimes the telephone company will state that the number is working but not whether it is assigned to a residence. Different estimates of the response rate will result if we assume that all, none, or some fraction of these numbers that were never answered were residences. In table 3, we present the range of screening response rates that would result from varying the assumption. We have used the midpoint to get the estimated overall response rate.

When interviewers are calling the first number in a primary sampling unit, we instruct them to call during business hours, since that will identify many non-residential primary sampling units quickly. Second and subsequent telephone numbers within a primary sampling unit should first be called in the

evenings and at other times when someone is likely to be home at a residence. On the average, six calls are made to complete the screening of each residence.

The telephone interview includes several questions needed to assure that the sample of individuals will be a probability sample of individuals in households with telephones in the designated geographic area. First, the interviewer confirms that the number reached is the number dialed and then determines whether the telephone belongs to a residence or a business. Since the telephone exchanges can encompass areas outside the designated one, the interviewer asks if the household is in the designated area and terminates

the call if it is outside. The interviewer also asks if the household has two or more separate telephone numbers. Because multiple residential telephone numbers increase the probability of a household being drawn by random digit dialing, such households should be subsampled in correspondence with the number of separate residential lines. In practice, this is a negligible problem.

RESULTS

Four epidemiologic studies (2–5) in which we used telephone household sampling are summarized in table 3. In all of these studies, households were used as a basis to draw a stratified random sample of individuals. Therefore, the household

TABLE 3
Random digit dialing in four case-control studies in the United States

	Study subject (ref. no.)			
	Bladder cancer (2)	T-cell lymphoma (3)	Thyroid cancer (4)	Cervical cancer (5)
Population	8 areas*, males and females aged 21–64 years	Philadelphia, other US, males and females aged 25–74 years	Connecticut, males and females aged 18–64 years	5 areas†, females aged 20–69 years
Year(s)	1979	1982	1981	1982–1984
Total households responding	21,447	1,231	1,007	11,864
Household-screening response rate (%)§	88–89	87–91	77–83	84–87
Sampling strata	Area, sex, age (5-year age groups)	Sex, race, county, age (5-year age groups)	Sex, age (5-year age groups)	Race, telephone exchange, age (5-year age groups)
Total subjects responding	2,265	92	261	795
Subject interview response rate (%)	84	78	78	72
Type of interview	In person	Telephone	In person	In person
Estimated overall response rate (%)	74	69	62	61

* Eight areas of the bladder cancer study: Seattle, Atlanta, New Orleans, San Francisco, Iowa, Detroit, Connecticut, and New Jersey.

† Five areas of the cervical cancer study: Alabama, Miami, Philadelphia, Chicago, and Denver.

§ Conservative and liberal estimates, based on assuming all or none of the telephones not answered were residential.

^{||} Preliminary results: total subjects responding and response rates not final.

interviews included asking for the age and sex, and sometimes the race, of each adult household member. In the T-cell lymphoma study, the household-screening call was followed by a letter to the respondent, which was followed by a telephone interview. In the other studies, the interview was conducted in person in the subject's home. In all studies, the household-screening interview included asking for a household address and the full name of any subject selected. The interviews varied in length and complexity.

In sampling households and interviewing for the limited purpose of getting a household census and sometimes an address, our response rates ranged from 77 to 88 per cent, conservatively estimated. The response rates during the interview phase ranged from 72 to 84 per cent. The estimated overall rate ranged from 61 to 74 per cent. Differences reflected various influences, including geographic variations, in the time and energy we spent to persuade initial refusers, in our sophistication and experience, and in telephone interview scripts and procedures. In the bladder cancer study (2), exactly the same telephone sampling procedures and staff were used for eight areas of the country. (Other procedures and staff were used in Utah and New Mexico.) Variations in the

response rate for the bladder cancer study therefore reflect differences in the populations and sampling variation. The results, shown in table 4, show a high rate in Iowa, a very rural state, and rather little variation otherwise. Response rates were 93 per cent in Utah and 86 per cent in New Mexico, both rural states.

DISCUSSION

Samples of individuals drawn from telephone-based samples of households as generated by random digit dialing can approximate true random samples of the population, depending on telephone ownership rates and overall response rates. In the United States and within many subpopulations, telephone coverage is almost universal as seen in table 1, but in certain subpopulations, e.g., among very poor or transient groups, it may be low. The second determinant of the representativeness of the telephone-based sample, the overall response rate, varies somewhat with the demographic characteristics of the population studied. In our experience, it depends greatly on the effort put into attaining a high rate. Telephone response rates can often be competitive with the response rates that can be achieved with other methods.

Other sampling frames that can be

TABLE 4
Household response rates in the National Bladder Cancer Study (2), by area, and reason for nonresponse*

Area	Nonresponse				Response	Total	Response rate (%)‡
	Refusal	No answer	Non-English	Other†			
Seattle	113	14	7	2	1,491	1,627	92
Atlanta	186	29	3	5	1,547	1,770	87
New Orleans	114	22	6	6	809	957	85
San Francisco	450	135	86	32	4,162	4,865	86
Iowa	120	64	0	8	2,586	2,778	93
Detroit	203	43	9	3	2,110	2,368	89
Connecticut	293	65	33	3	2,736	3,130	87
New Jersey	706	160	111	27	6,006	7,010	86

* Excludes New Mexico and Utah where different staff were used.

† Vacation, illness, or deaf.

‡ Conservatively estimated by assuming all telephones not answered were residential.

used to select a sample of the general population include: lists of telephone numbers, lists of dwellings, voter registration lists, driver's license lists, town lists (i.e., of all residents), and specialized population rosters, e.g., the Health Care Financing Administration roster of older Americans.

Dwelling-based (or area-based) sampling, like telephone-based sampling, is multistage, leading first to households and then to individuals. The other options are single-stage, leading directly to individuals.

Town lists and specialized rosters, if complete, can provide very inexpensive sampling frames. Voter registration lists typically exclude about one third of the adult population (9), do not list age, often list old addresses, usually are not computerized, and are not routinely revised to remove people who have moved or died. Driver's license lists typically exclude 10–15 per cent of the population (depending upon age) (10), are often computerized, and have addresses that are about two years old, on the average.

Because most individuals do live in dwellings, dwelling-based samples can provide good approximations to random samples of individuals in a population if the response rate is high. Dwelling lists that have already been created may also be inexpensive to use, depending on the demographic characteristics of the sample sought, but they are usually very expensive to create. For this reason, they have been used infrequently in epidemiologic research.

For both telephone-based and dwelling-based sampling, the cost of identifying a control depends upon the number of

households that need to be contacted to locate the control. The restrictiveness of the sampling selection criteria and the demographic profile of the population determine the difficulty and therefore the cost of identifying controls. For example, older men are more expensive to locate than other people because there are fewer of them per household.

In short, if complete population rosters are unavailable and if the population to be sampled has the high rates of telephone ownership typical of much of the United States, telephone-based sampling can yield a nearly random sample of the individuals in a population, often at much less expense than can dwelling-based sampling. In our experience, random digit dialing can be a valuable part of many epidemiologic studies.

REFERENCES

1. Glasser GJ, Metzger GD. Random digit dialing as a method of telephone sampling. *J Marketing Res* 1972;9:59-64.
2. Hartge P, Cahill J, West D, et al. Design and methods in a multi-center case-control interview study. *Am J Public Health* 1984;74:52-6.
3. Kantor AK, Greene M. Study of T-cell lymphoma. Unpublished manuscript.
4. Ron E, Kleinerman RA, Boice JD Jr., et al. A population based study of thyroid cancer in Connecticut. Unpublished manuscript.
5. Brinton LA. Study of invasive cervical cancer in the United States. Unpublished manuscript.
6. Waksberg J. Sampling methods for random digit dialing. *J Am Stat Assoc* 1978;73:40-6.
7. Mitofsky W. Sampling of telephone households. Unpublished CBS memorandum, 1970.
8. Harlow B, Hartge P. Telephone household screening and interviewing. Letter to the Editor. *Am J Epidemiol* 1983;117:632-3.
9. US Bureau of the Census. Voting and registration in the election of November 1980. (Population characteristics, Series P-20, No. 359.) Washington, DC: US GPO, 1980.
10. US Bureau of the Census. Statistical abstract of the United States: 1980. 101st ed. Washington, DC: US GPO, 1980.